Parallel PDE-Constrained Optimization: Antenna Identification in Hyperthermia Cancer Treatment Planning

Olaf Schenk
Computer Science Department, CS
University of Basel, Switzerland

Joint work with:
A. Sameh, M. Manguoglu (Purdue University), F. Curtis (New York University),
M. Christen, M.Sathe (University of Basel), A. Wächter (IBM Research Watson)

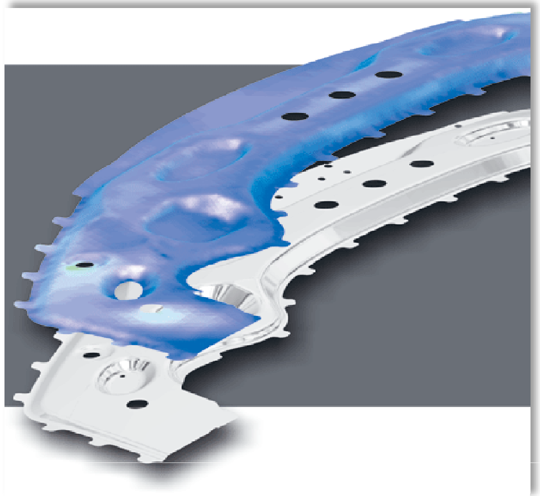
Overview

- Introduction Parallel Large-Scale Optimization
- Application: Hyperthermia Cancer Treatment Planning

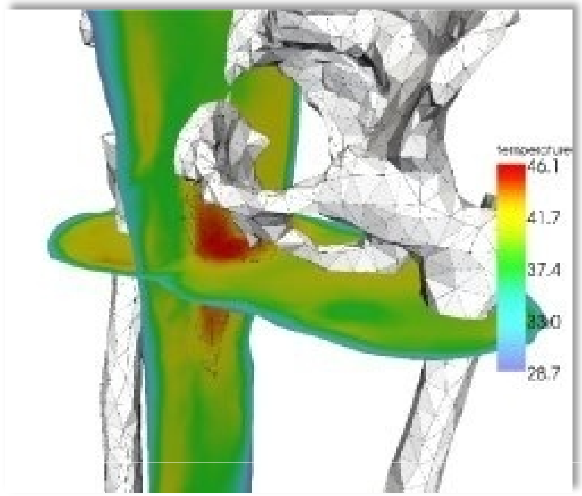
- How to solve it?
  - Highly Parallel Framework
  - Interior-Point Method: IPOPT (with IBM Research Labs)
  - Parallel Linear Equation Solver: PSPIKE (with U Purdue)
  - Preconditioner: Bipartite Graph Matching, Graph Partitioning

- Results
- Conclusion

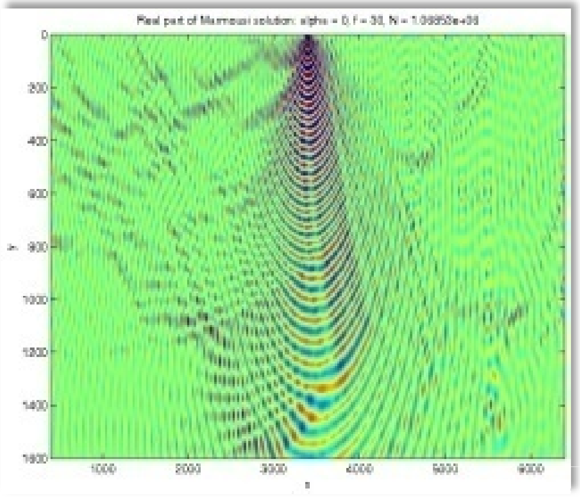
Large-Scale Nonlinear Optimization: Projects at U Basel



EU Project (BMW, AutoForm Engineering, U Basel)



SNF Project (ETH Zurich, U Basel, IBM Research, U Purdue)

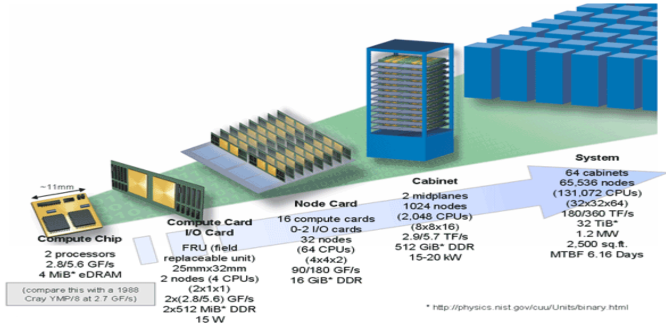


SNF Project (Shell, U Basel)

$$\boxed{\begin{array}{lll} \min\limits_{x\in\mathbb{R}^n} & f(x) & \\ \textbf{s.t.} & c(x) = 0, & c(x) : \mathbb{R}^n \to \mathbb{R}^m \\ & d(x) \geq 0, & d(x) : \mathbb{R}^n \to \mathbb{R}^q \end{array}}$$

Nonlinear Optimization: Parallel Architectures

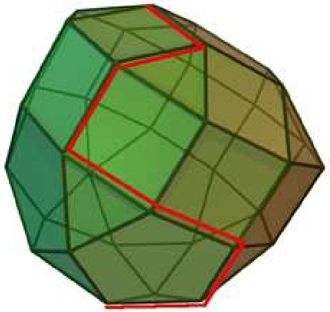
$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & f(x) \\ \textbf{s.t.} \quad & c(x) = 0 \\ & x \geq 0 \end{aligned}$$

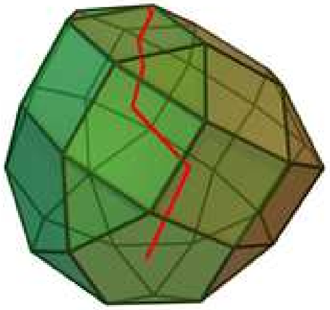


- Distributed-Memory Cluster at U Basel
  - 64 nodes each with eight Intel Xeon cores
  - Distributed–memory platform with Infiniband interconnection
- IBM BlueGene/L
  - 06/2009: position 3 in TOP500
  - 294‘912 cores with peak performance 825 Tflops
  - Our goal is to use 1‘000 processors efficiently

Nonlinear Optimization: General methods

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ \textbf{s.t.} \quad & c(x) = 0 \\ & x \geq 0 \end{aligned}$$





Nonlinear Optimization	Small-Scale (10^5 variables)	Large-Scale (10^6 to 10^9 variables)
Multicores (<16 cores)	- Simplex method (Linear Problems) - Randomized metaheuristics (Evolutionary Algorithms, Simulated Annealing, Ant Colony) - Derivative-free optimization - Interior-point optimization	- Interior-point optimization + Fast convergence - Need derivate information (Jacobian, Hessian matrices) - Matrices are indefinite and highly ill-conditioned
Manycores (~1‘000 cores)	———	- Interior-point optimization (will be addressed in this talk)

Nonlinear Optimization: Interior-Point Methods

- NLP Problem

$$\begin{array}{ll} \min\limits_{x \in \mathbb{R}^n} & f(x) \\ \textbf{s.t.} & c(x) = 0 \\ & x \geq 0 \end{array}$$

$\longrightarrow$

- Barrier Problem

$$\begin{array}{ll} \min\limits_{x \in \mathbb{R}^n} & \varphi_\mu(x) := f(x) - \mu \sum\limits_{i=1}^{n} \ln(x^{(i)}) \\ \textbf{s.t.} & c(x) = 0 \qquad \boxed{\mu \rightarrow 0} \end{array}$$

- Optimality Conditions

$$\begin{array}{rcl} \nabla \varphi_\mu(x) + \nabla c(x)\lambda & = & 0 \\ c(x) & = & 0 \\ (x & > & 0) \end{array}$$

$$H_k \approx \nabla^2_{xx} \mathcal{L}_\mu(x_k, \lambda_k)$$

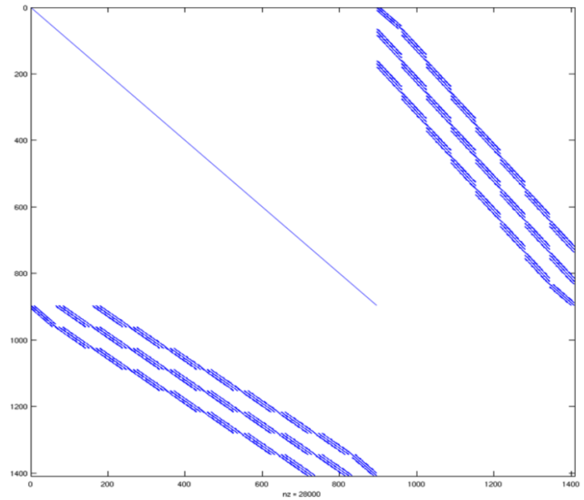
$$\mathcal{L}_\mu(x, \lambda) = \varphi_{\mu(x)} + c(x)^T \lambda$$

- Newton's Method for the Search Direction

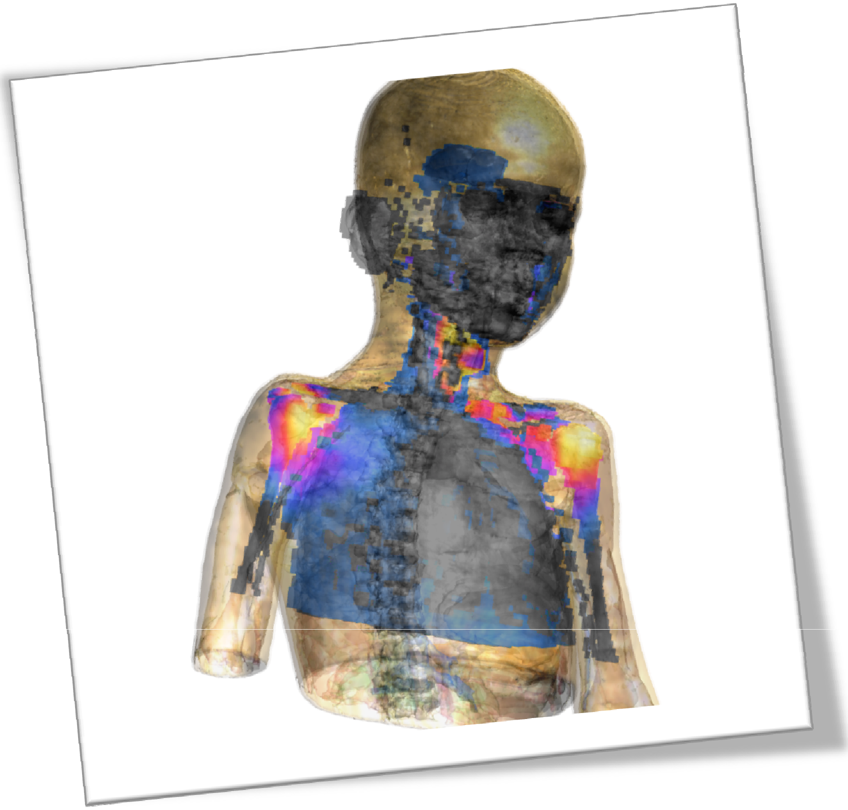
$$\begin{bmatrix} H_k & \nabla c(x_k) \\ \nabla c(x_k)^T & 0 \end{bmatrix} \begin{pmatrix} \Delta x_k \\ \Delta \lambda_k \end{pmatrix} = - \begin{pmatrix} \nabla \varphi_\mu(x_k) + \nabla c(x_k)\lambda_k \\ c(x_k) \end{pmatrix}$$

Parallel Nonlinear Optimization

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ \textbf{s.t.} \quad & c(x) = 0 \\ & x \geq 0 \end{aligned}$$



- **Parallel** components in **large-scale interior-point optimization**
  - +++ Computation of the objective function f(x)
  - +++ Generating the Hessian and Jacobian matrices
    - Taking sparsity of the matrices into account
  - - - - Distributed-memory sparse linear solver for Karush-Kuhn-Tucker systems
    - **Hybrid solver, Graph Partitioning+ Bipartite Matching**



APPLICATION

PDE–Constrained Optimization in Biomedical Hyperthermia Cancer Treatment Planning

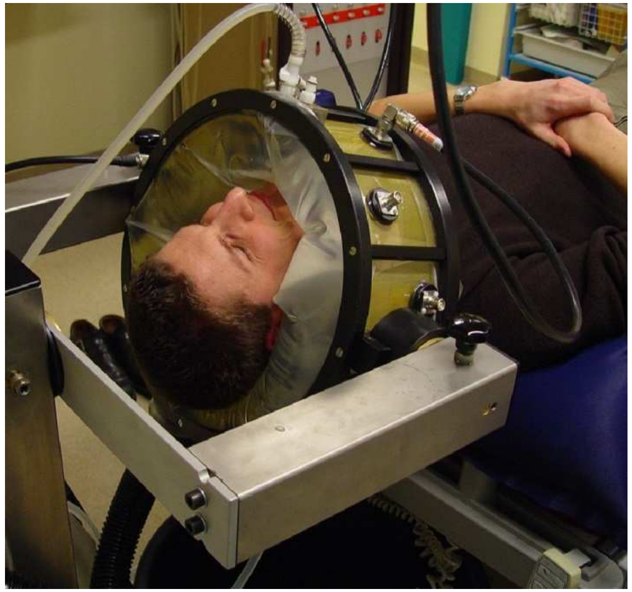
Hyperthermia Cancer Treatment

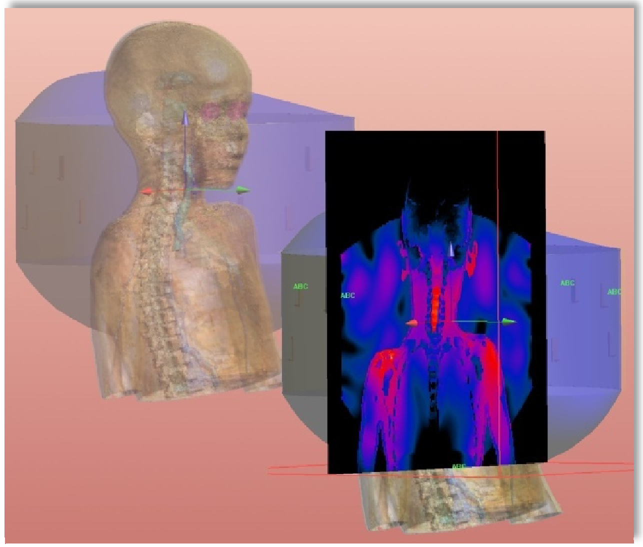
- **Hyperthermia** refers to various techniques of heat application to (parts of) the body
- **Hyperthermia cancer treatment** is usually used as an adjunct to other therapies (radiotherapy, chemotherapy)

- Apply heat to the tumor (41-45°C)

- The problem is typically formulated as a **PDE-constrained optimization** problem.

- Inequality constraints:
  State variables: temperature distribution **T**
  Control variable: electromagnetic antennas **u**

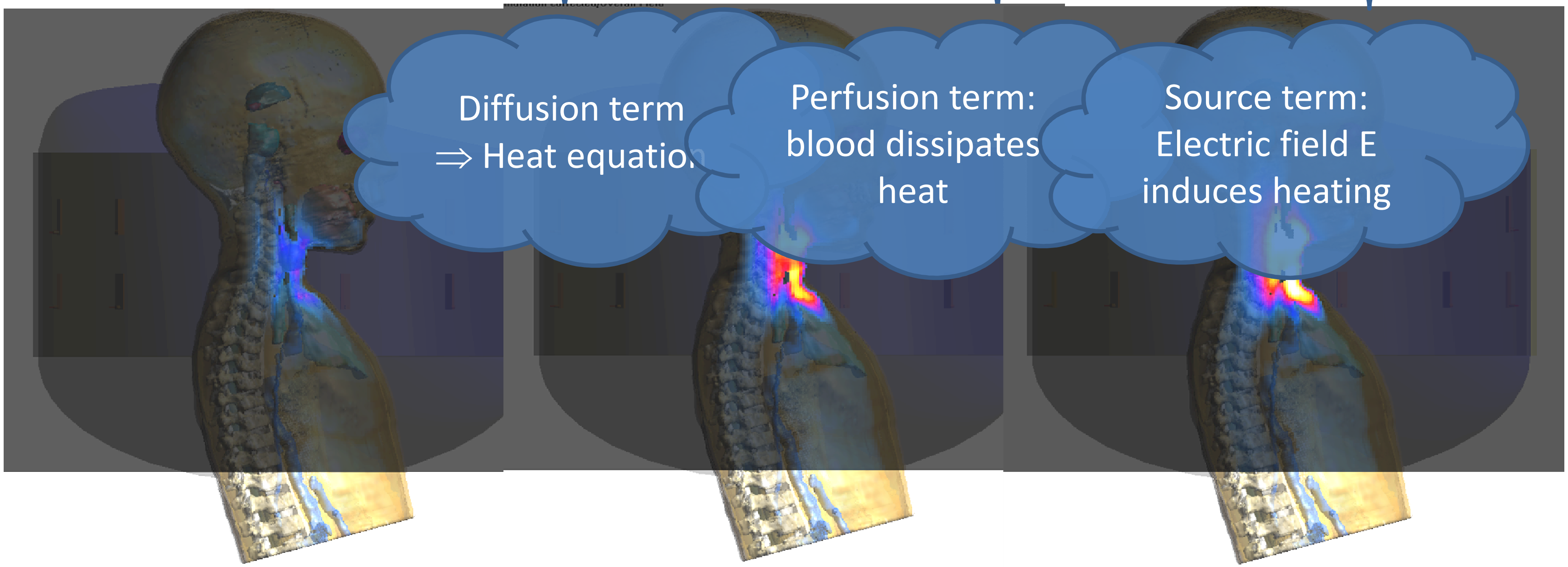




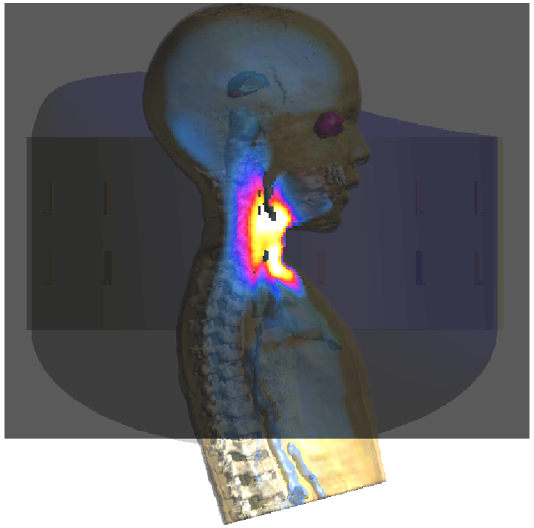
Hyperthermia — Forward problem

Penne's "Bioheat equation"

$$\rho C_\mathrm{p} \frac{\partial T}{\partial t} = \nabla \cdot (k \, \nabla T) - \rho_\mathrm{b} \omega_\mathrm{b} C_\mathrm{b} (T - T_\mathrm{b}) + \frac{\sigma}{2} \|\mathbf{E}\|^2$$



Hyperthermia — Inverse PDE-Constrained Problem



Minimize objective function

$$\min_{x \in \mathbb{R}^n} \quad f(x)$$

$$\min \int_{\Omega_{\text{tumor}}} (T_{\text{ther}} - T)^2 \, \mathrm{d}x + \alpha \, \|u\|$$

subject to the Pennes´ bioheat equation

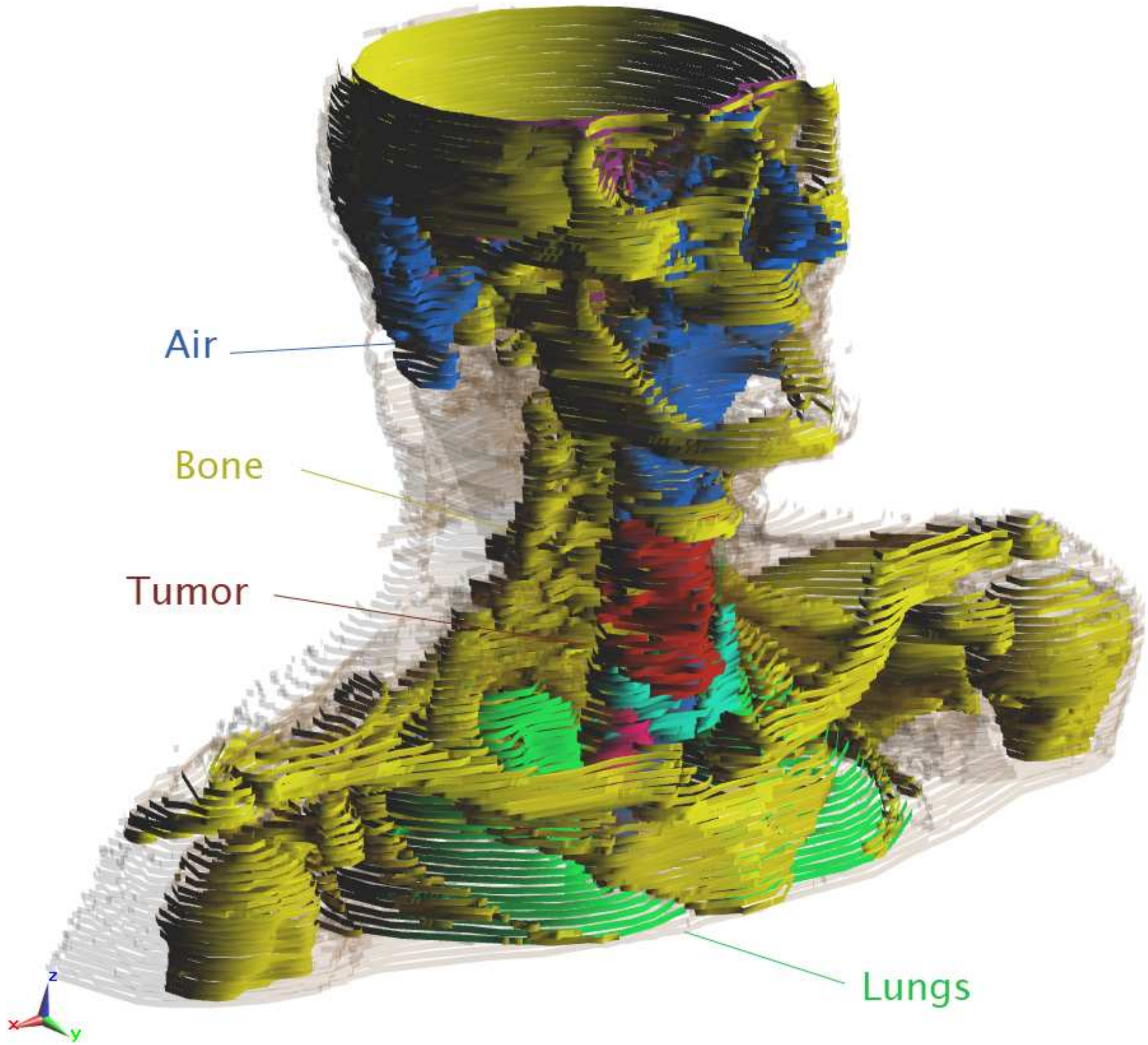
$$\textbf{s.t.} \quad c(x) = 0 \qquad x \geq 0$$

$$-\nabla \cdot (\kappa \nabla T) = -\rho \rho_{\mathrm{b}} \omega_{\mathrm{b}} C_{\mathrm{b}} (T - T_{\mathrm{b}}) + \frac{\sigma}{2} \left\| \sum_{i=1}^{n} u_i \mathbf{E}_i \right\|^2 \quad \text{in } \Omega$$

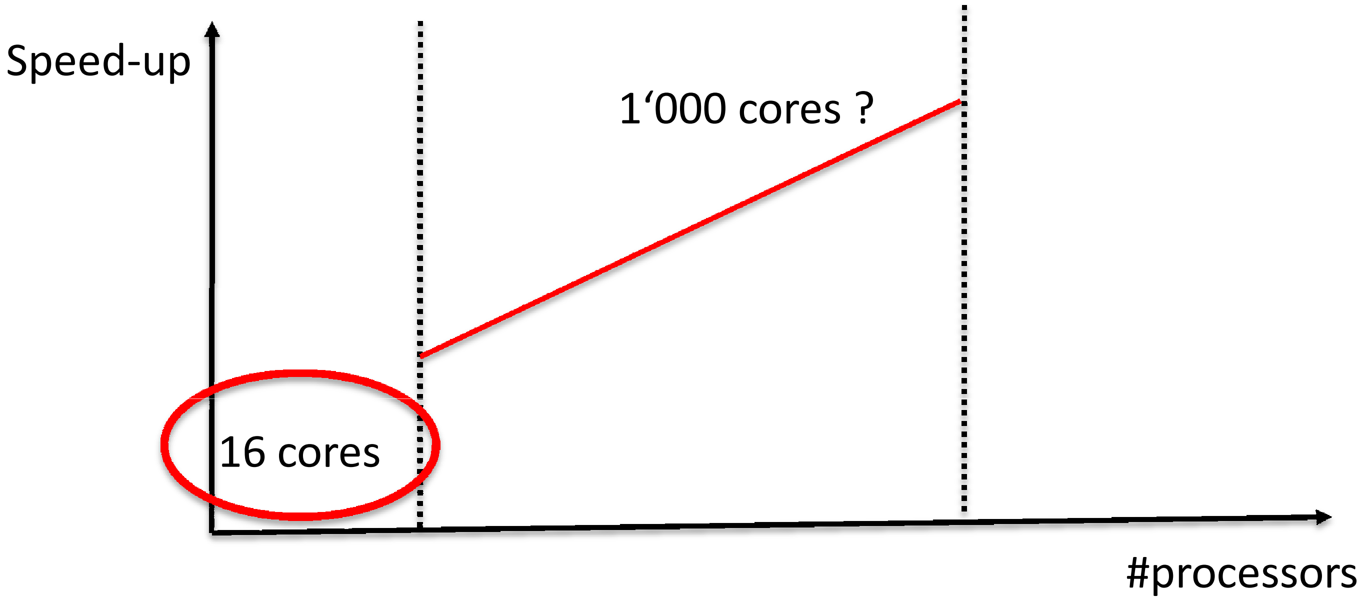
$$T = q \quad \text{on } \partial\Omega$$

$$T|_{\Omega \setminus \Omega_{\text{tumor}}} < T_{\text{limit}}, \qquad u_{\max} \geq u_i \geq u_{\min}$$

Optimization results from real hyperthermia patient data



Parallel Nonlinear Optimization: From small to large number of processors



- There is currently no state-of-the art interior-point optimization solver that scales up to more than 16 cores.
- Can we derive a **hybrid preconditioner** that scales better than sparse direct solvers and is also robust enough for interior-point optimization solvers?

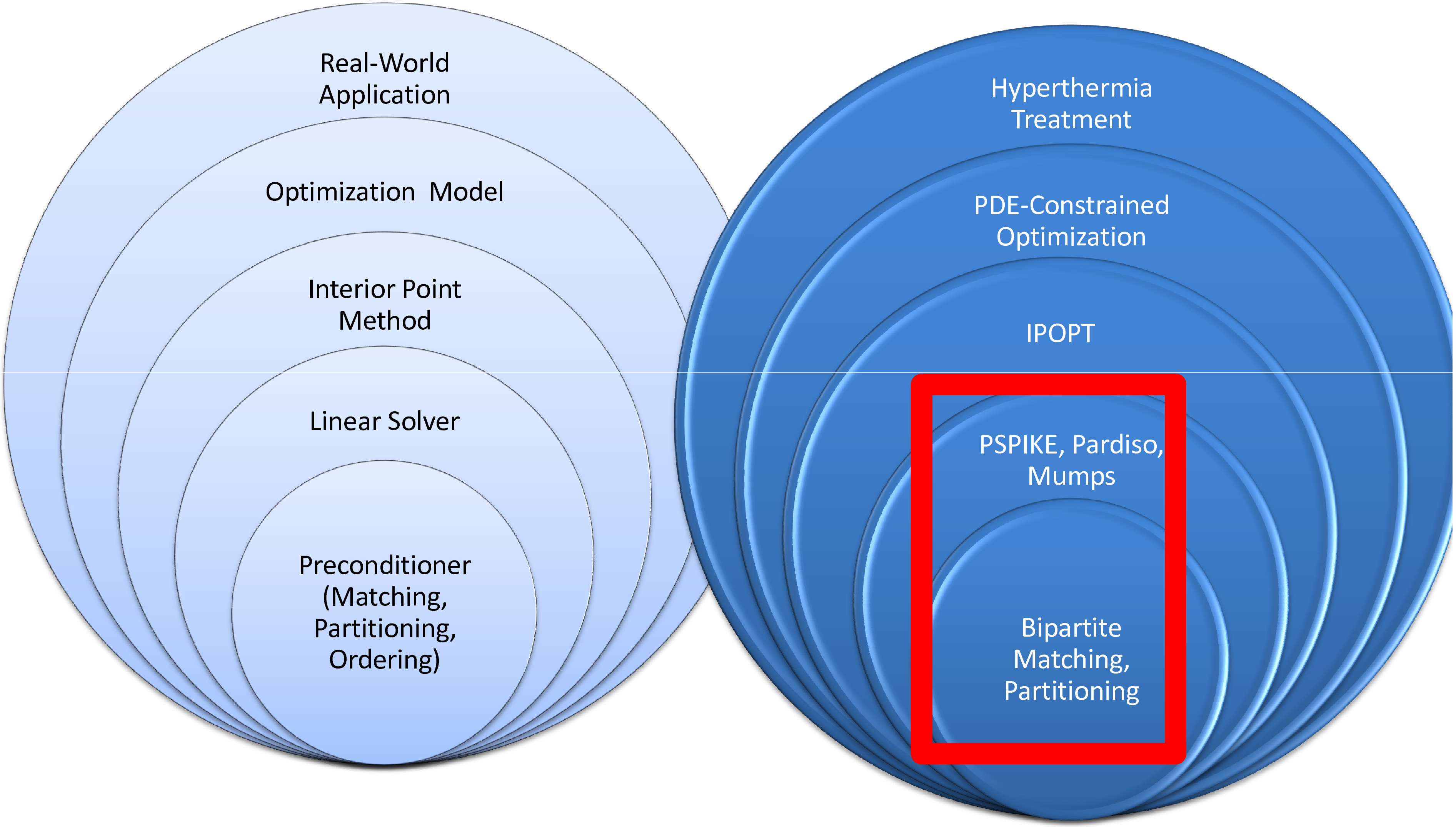
Challenges in PDE-Constrained Optimization

- **Discretization**
  - → Stencil computation
  - → Large number of constraints (10 to 100 millions)
  - → 23 Control variables (antennas)
- **Nonlinear and nonconvex**
  - – Objective function, blood perfusion -> Nonconvex PDE
- **How to solve it?**
  - – Highly scalable parallel framework
    - **Interior-point optimizer:** IPOPT (IBM Research, U Basel)
    - **Linear equation solver**: PSPIKE (U Basel, U Purdue)

Parallel Nonlinear Optimization: Main Ingredients

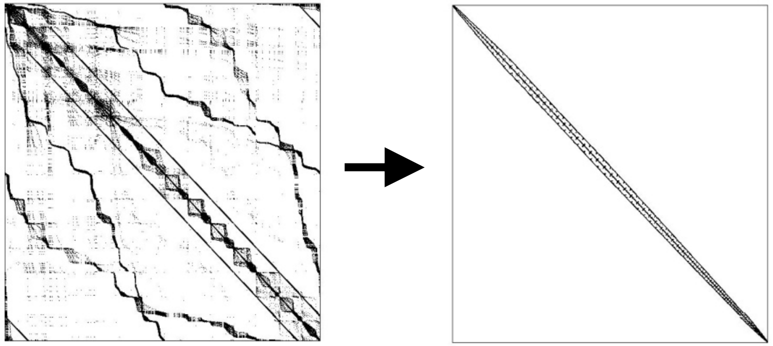
- **NLP optimizer IPOPT** (Wächter, S., IBM Research/U Basel):
  - Primal-dual interior-point solver (barrier method)
  - Line-Search algorithm for global convergence
  - Requires first and second derivatives of problem functions
  - Open-source C++, COIN-OR, > 5‘000 academic + industrial users
- **Linear equation solver PSPIKE** (Sameh, S., U Purdue/U Basel)
  - Hybrid **parallel** preconditioner
    - Linear ordering, graph matching, partitioning
  - Factorization in diagonal + spike matrices
  - Direct and iterative solver interactions

Overview: Parallel Optimization Framework



PSPIKE: Parallel preconditioner for KKT systems

- Many applications often produce large sparse linear systems
- Banded (or banded with low-rank perturbations) structure is often obtained after reordering

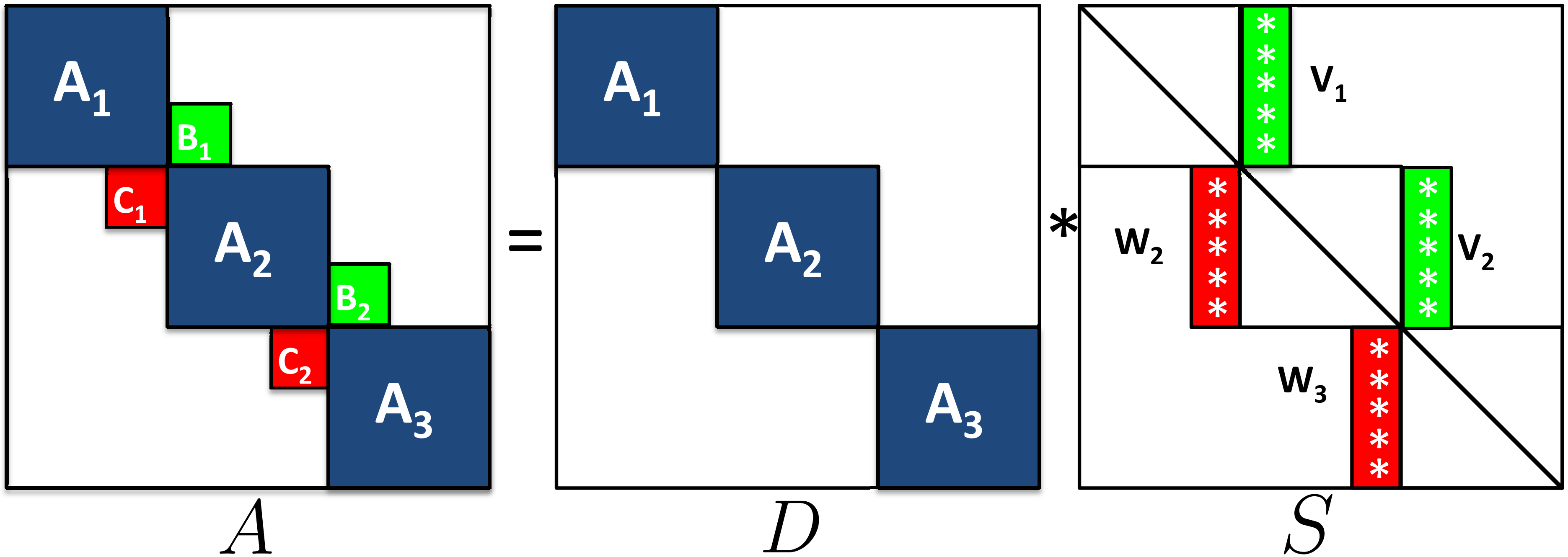


After variants of bipartite graph reorderings

- Almost all ILU preconditioners depend on reordering strategies that minimize the fill-in and maximize accuracy in the incomplete factorization stage. We propose to:
  - extract a sparse banded matrix, via a series of reordering and graph-partitioning strategies, to be used as a preconditioner.
  - make use of a novel PSPIKE scheme that is ideally suited for banded systems that are sparse within the band.

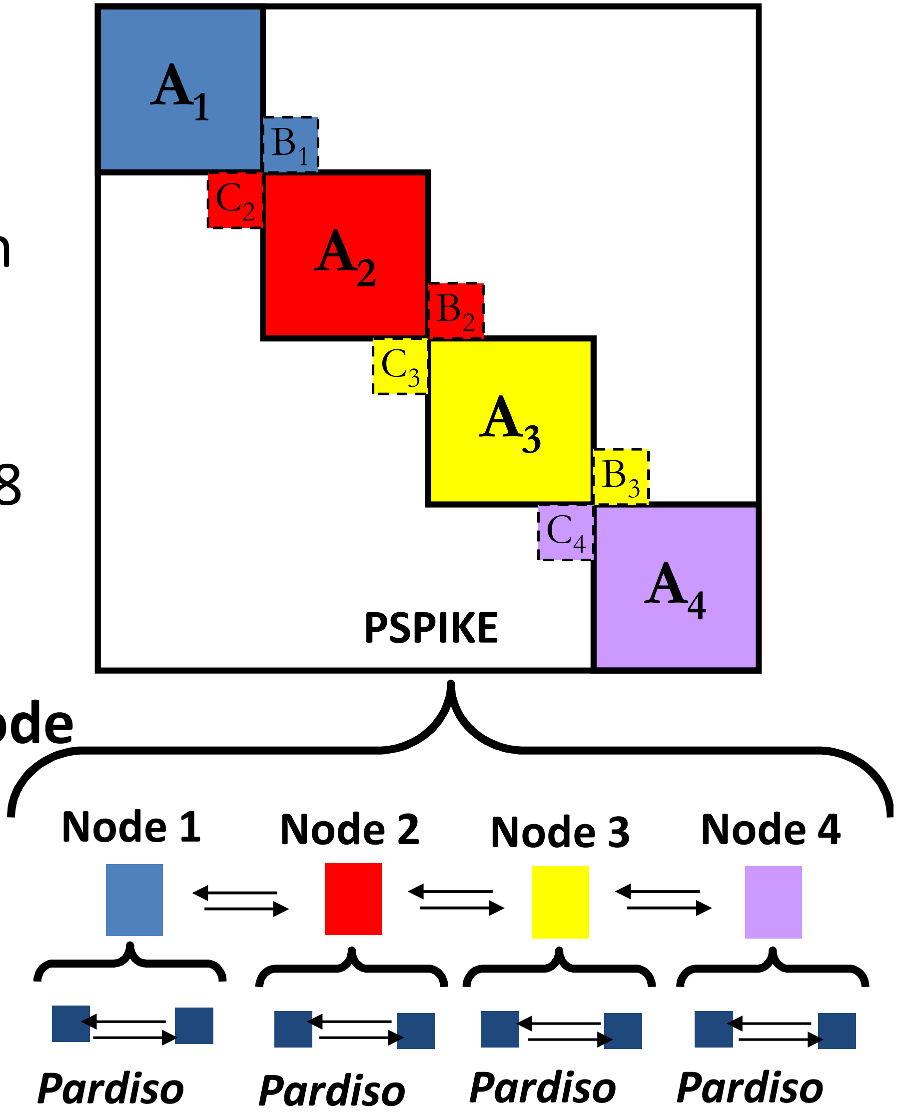
Basic PSPIKE scheme for sparse KKT systems

- **PARDISO** is an efficient direct linear solver for sparse matrices that scales well up to 8 to 16 cores on one SMP node (S., Gärtner, part of Intel's MKL)
- **SPIKE** is a parallel banded system solver proposed by A. Sameh (78), recently revisited (Polizzi, Sameh, part of Intel's Experimental Software What.if.intel.com)
- **PSPIKE** (S., Manguoglu, Sameh, ISC09) is a highly parallel robust preconditioner for symmetric indefinite matrices



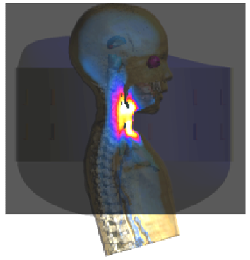
Computational Results

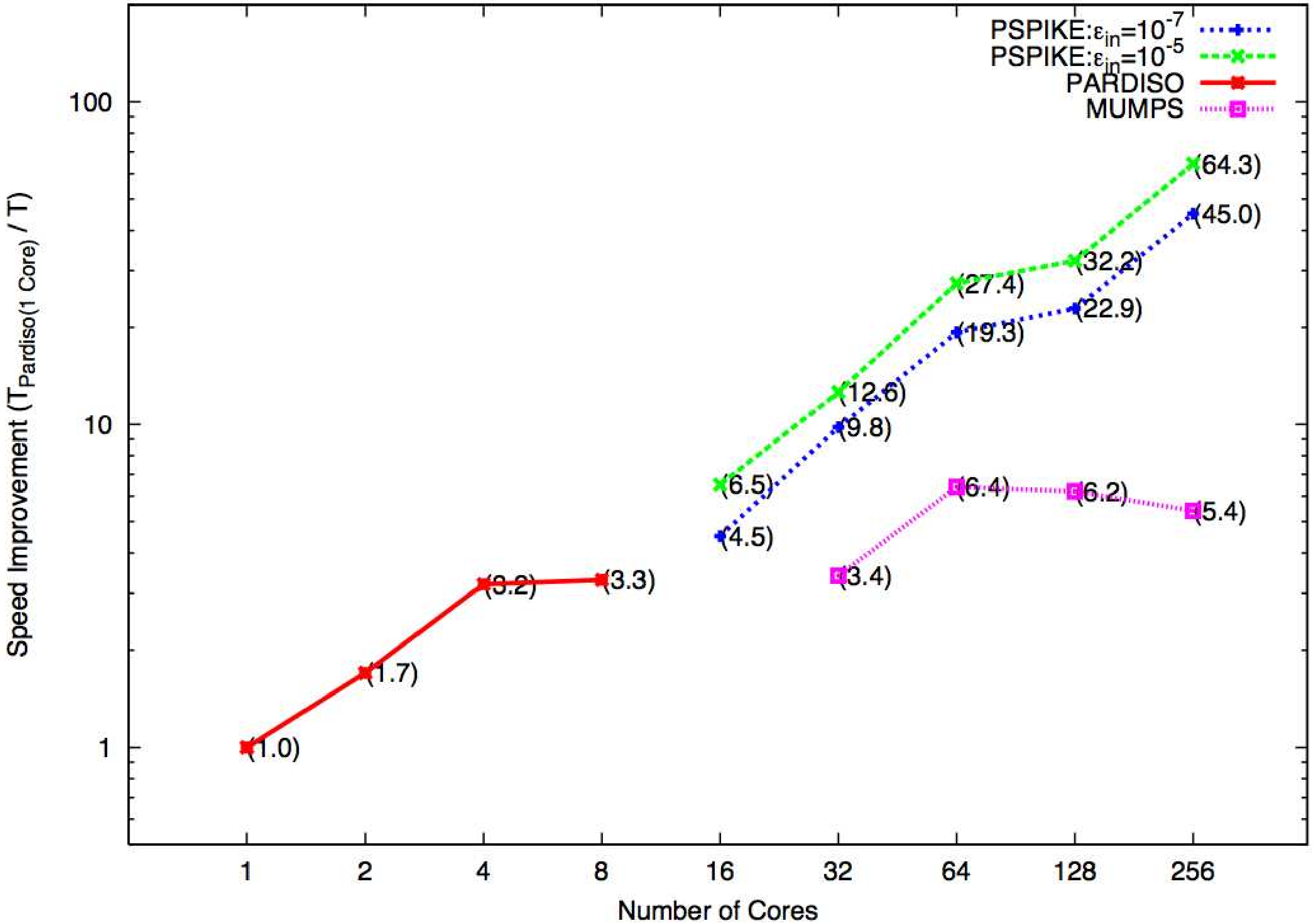
- **Parallel Architecture**
  - Distributed–memory platform with Infiniband interconnection
  - 64 nodes.
  - Each node is a dual quad-core Intel Hapertown processors (2,8 Ghz)



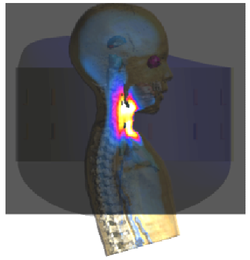
- **PSPIKE in mixed MPI-OpenMP mode**
  - MPI between nodes
  - OpenMP with the nodes
- All KKT matrices are scaled and ordered

PSPIKE – Scalability for one example matrix





Parallel Optimization – Timing results up to 512 cores



N	Threads	Nodes=1 (Direct)	Nodes=4 (Inexact)	Nodes=8 (Inexact)	Nodes=16 (Inexact)	Nodes=32 (Inexact)	Nodes=64 (Inexact)
75^3	1	32‘107 s. (1.0)	2‘016 s. (15.9)	1‘024 s. (31.1)	507 s. (63.2)	‡	‡
	4	10‘033 s. (3.2)	786 s. (40.8)	262 s. (122.5)	163 s. (201.0)	‡	‡
	8	7‘830 s. (4.1)	629 s. (51.0)	198 s. (162.1)	154 s. (208.2)	‡	‡
100^3	1	‡	7‘408 s.	3‘899 s.	1‘944 s.	934 s.	732 s.
	4	‡	2‘314 s.	1‘266 s.	664 s.	355 s.	345 s.
	8	‡	1‘763 s.	993 s.	462 s.	351 s.	301 s.
150^3	1	‡	60‘861 s.	33‘812 s.	17‘796 s.	9‘132 s.	6‘066 s.
	4	‡	20‘287 s.	10‘821 s.	5‘393 s.	4‘072 s.	3‘881 s.
	8	‡	14‘490 s.	7‘246 s.	4‘138 s.	1‘923 s.	1‘596 s.

Conclusion

- Highly Scalable Parallel Optimization Framework
  - General NLP method for large-scale optimization
  - It needs derivative information (Hessian, Jacobian)
  - Graph orderings (e.g. graph partitioning, bipartite weighted matching) are very important for convergence in numerical optimization
  - All steps are now highly parallelizable
- Nonlinear, nonconvex, and inequality constraints
- Scalabilty with up to 512 cores for a biomedical PDE-constrained optimization.
- General parallel optimizer -> can be applied to other applications.

Acknowledgments

- **Swiss National Science Foundation**, “Large-scale PDE-constrained optimization in hyperthermia cancer treatment planning”
- **IBM Faculty Award**, “Simulation, Modeling, and Optimization in Hyperthermia Cancer Treatment Planning”

Q & A

Thank you for your attention!

Questions?